RAW SEQUENCE LISTING DATE: 06/19/2001
PATENT APPLICATION: US/09/525,041 TIME: 12:33:29

Input Set : A:\PF178D2_SubSeqList.txt
Output Set: N:\CRF3\06192001\I525041.raw


```
      3 <110> APPLICANT: Soppet et al.
      5 <120> TITLE OF INVENTION: Colon Specific Gene and Protein
      7 <130> FILE REFERENCE: PF178D2
      9 <140> CURRENT APPLICATION NUMBER: US 09/525,041
C--> 10 <141> CURRENT FILING DATE: 2001-06-04
     12 <150> PRIOR APPLICATION NUMBER: US 09/162,508
     13 <151> PRIOR FILING DATE: 1998-09-29
     15 <150> PRIOR APPLICATION NUMBER: US 08/468,413
     16 <151> PRIOR FILING DATE: 1995-06-06
     18 <160> NUMBER OF SEQ ID NOS: 6
     20 <170> SOFTWARE: PatentIn version 3.0
     22 <210> SEQ ID NO: 1
     23 <211> LENGTH: 1114
     24 <212> TYPE: DNA
     25 <213> ORGANISM: Homo sapiens
     27 <220> FEATURE:
     28 <221> NAME/KEY: CDS
     29 <222> LOCATION: (111)..(587)
     31 <400> SEQUENCE: 1
     32 gcacgaggcc aaacagattt gcagatcaag gagaacccag gagtttcaaa gaagcgctag     60
     34 taaggtctct gagatccttg cactagctac atcctcaggg taggaggaag atg gct       116
     35                                                        Met Ala
     36                                                        1
     38 tcc aga agc atg cgg ctg ctc cta ttg ctg agc tgc ctg gcc aaa aca      164
     39 Ser Arg Ser Met Arg Leu Leu Leu Leu Ser Cys Leu Ala Lys Thr
     40         5                  10                  15
     42 gga gtc ctg ggt gat atc atc atg aga ccc agc tgt gct cct gga tgg      212
     43 Gly Val Leu Gly Asp Ile Ile Met Arg Pro Ser Cys Ala Pro Gly Trp
     44     20                  25                  30
     46 ttt tac cac aag tcc aat tgc tat ggt tac ttc agg aag ctg agg aac      260
     47 Phe Tyr His Lys Ser Asn Cys Tyr Gly Tyr Phe Arg Lys Leu Arg Asn
     48 35                  40                  45                  50
     50 tgg tct gat gcc gag ctc gag tgt cag tct tac gga aac gga gcc cac      308
     51 Trp Ser Asp Ala Glu Leu Glu Cys Gln Ser Tyr Gly Asn Gly Ala His
     52                 55                  60                  65
     54 ctg gca tct atc ctg agt tta aag gaa gcc agc acc ata gca gag tac      356
     55 Leu Ala Ser Ile Leu Ser Leu Lys Glu Ala Ser Thr Ile Ala Glu Tyr
     56             70                  75                  80
     58 ata agt ggc tat cag aga agc cag ccg ata tgg att ggc ctg cac gac      404
     59 Ile Ser Gly Tyr Gln Arg Ser Gln Pro Ile Trp Ile Gly Leu His Asp
     60         85                  90                  95
     62 cca cag aag agg cag cag tgg cag tgg att gat ggg gcc atg tat ctg      452
     63 Pro Gln Lys Arg Gln Gln Trp Gln Trp Ile Asp Gly Ala Met Tyr Leu
     64     100                 105                 110
     66 tac aga tcc tgg tct ggc aag tcc atg ggt ggg aac aag cac tgt gct      500
     67 Tyr Arg Ser Trp Ser Gly Lys Ser Met Gly Gly Asn Lys His Cys Ala
     68 115                 120                 125                 130
```

RAW SEQUENCE LISTING DATE: 06/19/2001
PATENT APPLICATION: US/09/525,041 TIME: 12:33:29

Input Set : A:\PF178D2_SubSeqList.txt
Output Set: N:\CRF3\06192001\I525041.raw

```
70 gag atg agc tcc aat aac aac ttt tta act tgg agc agc aac gaa tgc      548
71 Glu Met Ser Ser Asn Asn Asn Phe Leu Thr Trp Ser Ser Asn Glu Cys
72                 135                 140                 145
74 aac aag cgc caa cac ttc ctg tgc aag tac cga cca tag agcaagaatc      597
75 Asn Lys Arg Gln His Phe Leu Cys Lys Tyr Arg Pro
76             150                 155
78 aagattctgc taactcctgc acagccccgt cctcttcctt tctgctagcc tggctaaatc    657
80 tgctcattat ttcagagggg aaacctagca aactaagagt gataagggcc ctactacact    717
82 ggctttttta ggcttagaga cagaaacttt agcattggcc cagtagtggc ttctagctct    777
84 aaatgtttgc cccgccatcc ctttccacag tatccttctt ccctcctccc ctgtctctgg    837
86 ctgtctcgag cagtctagaa gagtgcatct ccagcctatg aaacagctgg gtctttggcc    897
88 ataagaagta aagatttgaa gacagaagga agaaactcag gagtaagctt ctagacccct    957
90 tcagcttcta cacccttctg ccctctctcc attgcctgca ccccacccca gccactcaac   1017
92 tcctgcttgt ttttcctttg gccataggaa ggtttaccag tagaatcctt gctaggttga   1077
94 tgtgggccat acattccttt aataaaccat tgtgtac                           1114
97 <210> SEQ ID NO: 2
98 <211> LENGTH: 158
99 <212> TYPE: PRT
100 <213> ORGANISM: Homo sapiens
102 <400> SEQUENCE: 2
104 Met Ala Ser Arg Ser Met Arg Leu Leu Leu Leu Leu Ser Cys Leu Ala
105 1               5                  10                  15
108 Lys Thr Gly Val Leu Gly Asp Ile Ile Met Arg Pro Ser Cys Ala Pro
109             20                  25                  30
112 Gly Trp Phe Tyr His Lys Ser Asn Cys Tyr Gly Tyr Phe Arg Lys Leu
113         35                  40                  45
116 Arg Asn Trp Ser Asp Ala Glu Leu Glu Cys Gln Ser Tyr Gly Asn Gly
117     50                  55                  60
120 Ala His Leu Ala Ser Ile Leu Ser Leu Lys Glu Ala Ser Thr Ile Ala
121 65                  70                  75                  80
124 Glu Tyr Ile Ser Gly Tyr Gln Arg Ser Gln Pro Ile Trp Ile Gly Leu
125                 85                  90                  95
128 His Asp Pro Gln Lys Arg Gln Gln Trp Gln Trp Ile Asp Gly Ala Met
129             100                 105                 110
132 Tyr Leu Tyr Arg Ser Trp Ser Gly Lys Ser Met Gly Gly Asn Lys His
133         115                 120                 125
136 Cys Ala Glu Met Ser Ser Asn Asn Asn Phe Leu Thr Trp Ser Ser Asn
137     130                 135                 140
140 Glu Cys Asn Lys Arg Gln His Phe Leu Cys Lys Tyr Arg Pro
141 145                 150                 155
144 <210> SEQ ID NO: 3
145 <211> LENGTH: 26
146 <212> TYPE: DNA
147 <213> ORGANISM: Artificial Sequence
149 <220> FEATURE:
150 <223> OTHER INFORMATION: Contains a BamHI restriction enzyme site.
152 <400> SEQUENCE: 3
153 gcaggatcct ggcttccaga agcatg                                       26
156 <210> SEQ ID NO: 4
```

RAW SEQUENCE LISTING DATE: 06/19/2001
PATENT APPLICATION: US/09/525,041 TIME: 12:33:29

Input Set : A:\PF178D2_SubSeqList.txt
Output Set: N:\CRF3\06192001\I525041.raw

157 <211> LENGTH: 28
158 <212> TYPE: DNA
159 <213> ORGANISM: Artificial Sequence
161 <220> FEATURE:
162 <223> OTHER INFORMATION: Contains complementary sequences to an Asp718 restriction enzyme
163 site.
165 <400> SEQUENCE: 4
166 tacgggtacc ttgctctatg gtcggtac 28
169 <210> SEQ ID NO: 5
170 <211> LENGTH: 36
171 <212> TYPE: DNA
172 <213> ORGANISM: Artificial Sequence
174 <220> FEATURE:
175 <223> OTHER INFORMATION: Contains a BamHI restriction enzyme site followed by 6 nucleotide
176 s resembling an efficient signal for the initiation of translatio
177 n in eukaryotic cells.
179 <400> SEQUENCE: 5
180 atcgggatcc gccatcatgg cttccagaag catgcg 36
183 <210> SEQ ID NO: 6
184 <211> LENGTH: 28
185 <212> TYPE: DNA
186 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Contains complementary sequences to an Asp718 restriction enzyme
190 site.
192 <400> SEQUENCE: 6
193 tacgggtacc ttgctctatg gtcggtac 28

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/525,041

DATE: 06/19/2001
TIME: 12:33:30

Input Set : A:\PF178D2_SubSeqList.txt
Output Set: N:\CRF3\06192001\I525041.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date